

Estimating Party Positions on Immigration: Assessing the Reliability and Validity of Different Methods

Didier Ruedin

didier.ruedin@unine.ch

Laura Morales

POST-PRINT

This is the final draft *after* refereeing

Published as: Ruedin, Didier, and Laura Morales. 2017. "Estimating Party Positions on Immigration: Assessing the Reliability and Validity of Different Methods." *Party Politics* Online First. doi:10.1177/1354068817713122.

<http://journals.sagepub.com/doi/full/10.1177/1354068817713122>

Supplementary material: <https://osf.io/bj27x/>

Abstract

We provide a systematic assessment of various methods to position political parties on immigration, a policy domain that does not necessarily overlap with left–right and is characterized by varying salience and issue complexity. Manual and automated coding methods drawing on 283 party manifestos are compared – manual sentence-by-sentence coding using a conventional codebook, manual coding using checklists, automated coding using Wordscores, Wordfish and keywords. We also use expert surveys and the Comparative Manifesto Project (CMP), covering the main parties in Austria, Belgium, France, Ireland, the Netherlands, Spain, Switzerland and the United Kingdom, between 1993 and 2013. We find high levels of consistency between expert positioning, manual sentence-by-sentence coding and manual checklist coding and poor or inconsistent results with the CMP, Wordscores, Wordfish and the dictionary approach. An often-neglected method – manual coding using checklists – offers resource efficiency with no loss in validity or reliability.

Estimating Party Positions on Immigration: Assessing the Reliability and Validity of Different Methods

Abstract

We provide a systematic assessment of various methods to position political parties on immigration, a policy domain that does not necessarily overlap with left-right and is characterized by varying salience and issue complexity. Manual and automated coding methods drawing on 283 party manifestos are compared – manual sentence-by-sentence coding using a conventional codebook, manual coding using checklists, automated coding using Wordscores, Wordfish, and keywords. We also use expert surveys and the Comparative Manifesto Project (CMP), covering the main parties in Austria, Belgium, France, Ireland, the Netherlands, Spain, Switzerland, and the UK, between 1993 and 2013. We find high levels of consistency between expert positioning, manual sentence-by-sentence coding, and manual checklist coding; and poor or inconsistent results with the CMP, Wordscores, Wordfish, and the dictionary approach. An often-neglected method – manual coding using checklists – offers resource efficiency with no loss in validity or reliability.

Keywords: party positions, party manifestos, immigration, Europe, position estimation

Introduction

For many questions in political science it is important to accurately measure the policy positions parties take on various issues, including problems of political representation and party competition. While the positions on broad ideological ‘bundles’ like the left-right and the authoritarian-libertarian dimension attract most attention, often specific issue domains are of interest. Here, we focus on immigration and the integration of migrants and their descendants. Although many policy fields characterize the ideological positions of parties, the management of immigration and diversity has become increasingly important in determining government formation and dynamics of party competition, and can affect key decisions like staying in the European Union – consider the 2016 referendum in the United Kingdom. Arguably, positions on immigration are a key component of new cleavages that do not always fully overlap with the left-right divide (Kriesi et al., 2006, 2012). Yet, we lack valid and reliable estimates of party positions on immigration across countries and over time.

There are various methods to capture the positions of political parties on specific issues or dimensions of political competition, and there is a considerable methodological debate around the validity and reliability of each of these methods (see special issue edited by Marks et al., 2007). Data from the *Comparative Manifesto Project* (CMP) and the *Chapel Hill Expert Survey* (CHES) are commonly used, often because they are available and established (McDonald and Budge, 2014). Sometimes, however, these data do not fit the theoretical or empirical concerns of the researchers (Laver, 2014). For instance, the coding scheme of the

CMP has bundled immigration issues with other issues and is ill-suited to place parties on immigration over time.¹ In such instances, new measurement approaches are needed, and party manifestos are a useful source because they signal the position that the party – as an organization – has taken to compete for votes.²

We examine the performance of several methods for measuring party positions on immigration and immigrant integration from manifestos. A specific challenge is that occasionally the relevant passages in manifestos are short, and the issue is not salient. We compare a wider range of approaches than previous studies – manual sentence-by-sentence coding, ‘checklist’ coding, pooled expert surveys, Wordscores, Wordfish, and an automated dictionary approach –, and go beyond correlations to examine the variance of estimates and *how* positions are associated. The range of methods examined in this article follows calls for (convergence) validity checks (Dinas and Gemenis, 2010; Grimmer and Stewart, 2013).

We do not seek to crown any particular approach as a victor, and demonstrate that all methods tested tend to pick up the same underlying construct – positions on immigration. Manual sentence-by-sentence coding, ‘checklist’ coding, and pooled expert surveys are consistently

1 See exact wording in supplement S1 [[ADD LINK TO ONLINE SUPPLEMENT](#)]. Codes 601 and 602 refer to national ways of life, not directly immigration. Codes 607 and 608 include considerations of pillarization, thus include ‘internal’ or ‘autochthonous’ minorities. Code 705 makes no distinction between immigrants, national minorities, or minorities such as homosexuals and the disabled. In the 5th edition of 2014 sub-codes were introduced differentiating immigrants specifically in sub-codes 607.2, 607.3, 608.2, 608.3.

2 For Britain and Sweden, Odmalm (2012) coded party positions on immigration (in general), labour immigration, asylum seekers and refugees, family reunification, unaccompanied minors, student migration, and retirement migration.

highly correlated. Sometimes, Wordscores, Wordfish, and the dictionary approach are also closely associated with these methods, but not universally so. The less known ‘checklist’ method is both efficient and positions parties congruent with the established sentence-by-sentence coding and expert surveys. We do not argue that ‘checklists’ are universally preferable over other methods, but they emerge as a cost-effective and valid method worth considering for future research applications.

Obtaining positions from manifestos: framework and debates

Political parties are complex organizations with several faces – on the ground, in central office, and in public office (Katz and Mair 1992, 1994) – that can all take positions on relevant political issues, sometimes contradicting each other. Parties rarely have a single, coherent and unequivocal position on any issue or policy area, which makes it impossible to find their ‘true’ position. Political scientists instead try to capture instances where one of these faces reveals where the party stands. The production of manifestos at the time of elections is one such instance. Party manifestos are available in most political contexts and can be analysed retrospectively. Party manifestos are available in most political contexts and can be analysed retrospectively. Written in the context of competitive elections, electoral manifestos outline the official and collectively agreed policy preferences and proposals that the party puts forward to the electorate. As party manifestos are available in most political contexts and can be analysed retrospectively, they are a convenient and sufficiently valid source of the revealed

positions of parties. We focus on approaches based on manifestos as political texts, acknowledging that there are other methods (see Laver, 2014).³

Using electoral manifestos as the source of party positions comes with conceptual, theoretical, and methodological challenges. Conceptually, what constitutes an electoral manifesto varies across countries, parties, and over time. Similarly, who decides the content of the manifesto and who drafts it also varies across parties. Hence, manifestos reflect the positions of a complex and varying set of actors within the party organization. Linking theoretical and methodological challenges, what constitutes a position and what is the issue space on which to locate such a position are not settled matters. First, manifestos treat different policy areas at varying length, and sometimes parties choose to remain silent on issues that place them in a competitive disadvantage (Budge and Farlie, 1983). This may lead to the incorrect conclusion that the party has no position on the issue, or that the issue is not salient for the party. Silence (or short references) might also mask internal disagreement on the issue, and the inability to reach a common position. Any method deriving positions from manifestos alone will face problems to interpret silence correctly.

Second, there is no agreement on how the competitive issue space in which parties position themselves – through party manifestos and public statements – is configured. Some primarily

3 Alternative methods include aggregated positions from mass surveys to determine how voters perceive party positions, but they are problematic to capture the positions verbalised by parties. Data on the behaviour of members of parliament, roll call data, legislative debate analysis, as well as media and claims-making analysis offer different approaches to positions verbalised by the party in ‘public office’ and in ‘central office’.

view it as competition driven by the minimization of the distance between the (median) voter and the parties along a continuum (Black, 1948; Enelow and Hinich, 1984); others as competition driven primarily by direction and valence, whereby salience and perceived competence determine vote choice and party behaviour (Budge and Farlie, 1983; Rabinowitz and Macdonald, 1989; Petrocik, 1996). These different conceptions of how voters choose among parties, and how parties compete for votes have consequences for how we extract positions from electoral manifestos.

In the case of immigration and the integration of migrants and their descendants, the different conceptualizations of the political space either mean positioning parties on a liberal/pro-immigration vs restrictive/anti-immigration continuum, or also taking into account the emphasis or salience parties give to immigration. Furthermore, competition around immigration policies can be perceived along a single continuum or in a multi-dimensional space, given that immigration and integration policies bundle a complex set of – sometimes contradictory – sub-issues. We argue that an optimal method for deriving positions on immigration gives researchers many choices as to which model of party competition is assumed, rather than imposing a single model embedded in the data collection. In the following, we discuss different methods that draw on party manifestos to determine party positions on immigration, and compare them to expert surveys – an established alternative source of party positions.

Manual approaches: Sentence-by-sentence coding vs. coding the manifesto as a whole

Party manifestos are usually long enough to provide detailed information about multiple issues or dimensions. Among the examples of manual coding, the work of the CMP stands out for its breadth of coverage (e.g. Budge et al., 2001). Over the years, the CMP has consistently applied a set of codes to quasi-sentences. Some scholars have criticized the failure to expand the coding scheme to emerging issues like immigration, the use of only one coder per manifesto, the dominance of the salience approach to party competition, the heterogeneity of the documents used as texts for coding, or the fluctuations in the positions of parties in successive elections arguably reflecting what parties choose to emphasise rather than real differences in policy positions (Benoit and Laver, 2006).

In many research projects coding full manifestos (quasi-)sentence-by-(quasi-)sentence is not an option due to costs and reliability problems if not enough is invested in training coders (Lowe and Benoit, 2013). Crowd-sourced coding can reduce cost and deliver valid and reproducible results for broad positions, if the coding scheme is simple (Benoit et al., 2016). Yet, this may not be suitable for studies focusing on nuances of position-taking or the various sub-issues and frames addressed by parties.

An alternative approach is coding manifestos as a whole. Political texts are treated as data, but coders use their own judgements as to which sections constitute evidence for a certain position. Expert coders read a manifesto and use a codebook to assign an overall position in a

number of policy or ideological domains. Harmel et al. (1995) used this approach in the Party Change Project (PCP);⁴ Gudbrandsen (2010) applied it to positions on refugee immigration in Norway coding three broad scores – restrictive, liberal and neutral (or no statements) –, but the approach is not in wider use. This ‘checklist’ approach is much less time-consuming than coding individual (quasi-)sentences and provides a global consideration of the policy positions in a manifesto, reducing the scope for random fluctuations due to writing styles or personal emphases of the manifesto author(s). By relying on a global assessment, however, the checklist approach is sensitive to coder biases, as prior information about the party may interfere more than when coders examine shorter (quasi-)sentences.

Automated approaches: Word frequencies, keyword matches, Wordscores, and Wordfish

Automated approaches draw on the relative frequency of words in manifestos to identify party positions, disregarding word order and the context in which a given word occurs. The underlying models of language are inaccurate (Grimmer and Stewart, 2013), but they may work empirically (Lowe and Benoit, 2013). Party positions are derived from the fact that parties emphasize different issues in their manifestos, and use different frames when treating the same issues. This leads to different words being used, reflected in relative word frequencies: words have strong directional associations – e.g. the term *tax* is almost exclusive

4 The codebook includes ‘minority’ rights (#16), but no distinction is made between racial, linguistic, and regional minorities. Code #18 covers immigration policy, focusing on the entry of immigrants. Covering 1950–1990, the data are not suitable for the analysis of more recent developments.

to parties on the right (Laver et al., 2003). Validation is essential for automated approaches (Grimmer and Stewart, 2013).

A different approach relying on word frequencies uses a dictionary of keywords. By drawing on in-depth knowledge of the language used in manifestos and the topic at hand (or any other suitable method), researchers draw up a list of words – usually word stems – that are matched by the computer. Each word can be assigned a score, such as -1 for negative towards immigrants, or +1 for positive towards immigrants. The choice of keywords and the assignment of scores constitute the codebook. The dictionary needs to be precise, as no human coders are involved in judging synonyms or spot false positives. Party positions are calculated on the relative occurrence of positive and negative keywords.

Wordscores reduces human input (and potential bias) further by doing away with the dictionary. Instead, reference texts with known positions – like those drawn from expert surveys – need to be specified. A problem particular to Wordscores is that it often appears less reliable for extreme positions (Lowe, 2008). Depending on the research question, the exact positions at the edges might be of crucial interest. In the case of immigration it seems difficult to find clear *pro*-migration stances as reference texts, as parties with more immigrant-friendly policies tend to include them as part of wider concerns for equality and diversity.

Wordfish uses a model to estimate policy positions from all the manifestos provided. Two manifestos need to be specified to identify the direction of the policy domain, but no reference

scores are required. In this sense, Wordfish further removes the input required from researchers. Like Wordscores, Wordfish relies on the assumption that words in the manifestos primarily carry ideological meaning – as opposed to purely stylistic ones. If this assumption is not met, Wordfish may fail to identify policy positions (Grimmer and Stewart, 2013).

Expert surveys

Expert surveys are often used to position parties on ideological or policy dimensions, either on the left-right scale (Castles and Mair, 1984; Huber and Inglehart, 1995) or on multiple dimensions, and they are adapted to new conflict lines as they emerge (Benoit and Laver, 2006; Hooghe et al., 2010; Rohrschneider and Whitefield, 2012). Specific issue domains are usually covered in a limited and generic manner. For instance, distinctions between immigration flow management and immigrant integration policy are rarely covered. Until recently, very few expert surveys enquired about immigration at all. Benoit and Laver (2006) included one such item (#19), asking experts to place parties on a 1–20 scale: “Favours policies designed to help asylum seekers and immigrants... integrate into [country] society (1) / return to their country of origin (20)”. The CHES series added a comparable item in 2006 (#25) asking experts to place parties on a 0–10 scale where 0 means strong opposition to tough policy and 10 means strong preference for tough policy. This item is followed by another (#27) on the integration of immigrants and asylum seekers (0 = strongly favours multiculturalism, 10 = strongly favours assimilation).

The advantage of expert surveys is that they are cost efficient and authoritative, can cover positions on multiple domains, and do not require complex data processing. The assumption is that experts (usually political scientists specialized in party politics) know where the parties in their countries stand on several issues and dimensions of political competition. Including several experts per country will, additionally, provide estimates of uncertainty. Expert surveys may be fuzzy about whose positions are measured (party leaders, party activists, party supporters), sometimes fail to specify the criteria to place parties, whether the positions reflect preferences and intentions or behaviour (or a mix of both), or what the time period of reference is, and their timing may not coincide with elections (see summaries in Benoit and Laver, 2007; Budge, 2000; Marks et al., 2007; Ray, 2007). Nevertheless, Steenbergen and Marks (2007) show that expert surveys are reliable and often display remarkable agreement among experts. It is thought that experts are unable to position parties reliably in retrospect (but see Ruedin, 2013), and that experts tend to discount smaller changes in party positions (McDonald et al., 2007), which may limit their usefulness for analyses over-time.

The rationale for comparing methods for positioning parties on immigration

Researchers may have preferences as to how to measure party positions, but there is no gold standard or universally preferred way. All methods are subject to known limitations and unknown biases. Our goal is to assess how estimates of positions on immigration stemming from different methods – all drawing on party manifestos – compare, and how they compare to expert positions as an alternative established source. Assuming that all methods try to

capture the same underlying continuum or immigration policy dimension we examine *how* estimates from different methods correspond, providing a sense of their relative validity, and the advantages and disadvantages each presents to capture that latent dimension. We do not seek to assess methods in their general ability to estimate party positions as a whole, to crown a victor, or to replicate policy positions from a given method like expert surveys.

Of course the methods outlined above have been compared before (Benoit and Laver, 2007; Klüver, 2009; Rooduijn and Pauwels, 2011), but here we are interested in *how well* different methods correspond in the case of immigration, and how they compare across a number of countries. The cross-national study of this policy domain presents certain challenges that are common to other specific policy domains, which suggests that our findings may be of broader relevance beyond the narrow study of party positions on immigration.

First, the attention parties pay to immigration varies considerably across parties, years and countries. Second, the language and expressions used – specifically the sub-issues and frames related to immigration – can vary from one election to another. For example, immigration might be discussed primarily in connection to crime in one year/country and in connection to cultural accommodation of diversity in another. Third, when researchers are interested in deriving party positions on immigration (including multiple sub-issues and frames) across countries and over time, they have to deal with different languages.⁵ Fourth, resources are

⁵ The crowd-sourcing trials by Benoit et al. (2016) only used English-language coders. Replicating this in languages with fewer potential coders may be a challenge.

often limited in projects that use party positions as a predictor variable. Our analysis informs researchers facing such challenges and constraints.

Data and methods

We examine party manifestos in 8 West European countries (Austria, Belgium, France, Ireland, the Netherlands, Spain, Switzerland, and the United Kingdom), covering 43 elections between 1993 and 2013. Most of the data were collected in the context of an international project on the politicization of immigration, aiming to maximize the likely variance in the salience of and position around immigration across time and countries (Van der Brug et al., 2015). In each country, we identified relevant parties – usually parties that gained seats in the legislature, but in some instances anti-immigrant parties without parliamentary representation because of the topic at hand. Overall 283 party manifestos were coded and analysed in different ways (supplement S2).

For manual coding, coders were provided with written instructions, and underwent standard training. Only the statements in party manifestos directly relating to immigration and integration were coded (the immigration corpus). Relevant sentences and paragraphs were identified manually (supplement S3), heavily assisted by a dictionary of keywords (supplement S4) and the automated search functions in *Yoshikoder*. The keywords were derived from expert knowledge of the authors, and assisted by word frequency analysis of manifestos from anti-immigrant parties. Tests with alternative sources for the compilation of

the dictionary (e.g. mainstream parties) show that our original source does not bias the dictionary (supplement S5). The keywords were translated and back-translated in several rounds to ensure equivalence across languages, and tested on several manifestos in English, German, and Spanish. The selection of relevant sections eases manual coding, and is necessary for automated approaches to ascertain validity.

Manual sentence-by-sentence coding uses a conventional codebook applied to natural sentences (supplement S6). Often texts are divided into quasi-sentences – natural, or parts of a sentence with independent meaning –, but this additional effort is not necessary (Däubler et al., 2012). The positional question asks “what is the position toward the issue of immigration and civic integration?”. Possible answers range from “Strongly restrictive to migrants/ conservative/ pro-national residents/ mono-cultural” (-1) to “Strongly open to migrants/ progressive/ cosmopolitan/ multi-cultural” (+1) on a five-point scale. Examples were included to aid coding. 8,099 sentences were coded manually. Alternative ways of estimating party positions were employed: the mean of the positions assigned to each sentence in a manifesto, but also the median and interpolated median (Revelle, 2015). Given the ordinal nature of the data, we focus on interpolated medians in the presentation of the results. Because multiple sentences are coded per manifesto, we can quantify variance.

To code manifestos as a whole (checklist approach), a questionnaire using 19 statements was created (supplement S7). Coders were asked to read the entire manifesto corpus on immigration, and determine whether the party agrees or disagrees with each of the 19

statements in relation to the manifesto corpus as a whole (e.g. “There are too many refugees”). The remainder (neutral, not mentioned) was coded as zero. To minimize bias from external information, preferences, and prior knowledge, coders were asked to copy snippets from the manifestos as evidence. We could not identify significant coder biases (supplement S8). The mean of the coded positions is taken as the party position. With 19 categories, the use of the mean is unproblematic.⁶ Here we only focus on the overall position on immigration, but the data obtained for each of the 19 categories — or bundles of these — can be used to position parties on a number of sub-issues relating to immigration and integration.

Automated dictionary coding was implemented using *Yoshikoder*. The dictionary is the same used to identify parts on immigration, but words are assigned scores. Extensive testing and several rounds of translating and back-translating were used to refine the keywords and assign scores to the keywords (supplement S4), thus combining the use of traditional human coder keyword generation practices with automated procedures to produce a dictionary that was as exhaustive as possible while still operational across several languages. By only using sections of the manifesto known to be about immigration, scores could be assigned to some keywords that would otherwise be ambivalent. Party positions were calculated on the basis of scores assigned to keywords (positive matches minus negative matches divided by any match,

⁶ Because sub-issue focus can be country- or election-specific, we also calculated the mean of the positions on which any of the parties took a position (‘issue-space position’). This alternative operationalization accounts for the fact that in some elections/countries some aspects of immigration are not politicized: only sub-issues where at least one party took a positive or negative position were included. The two approaches correlate strongly ($r=0.97$).

including neutral matches). With no natural bounds to the generated scores, the minimum and maximum across all manifestos served as empirical anchors (supplement S9).

Wordscores and Wordfish were applied using *JFreq* and the *Austin* package in R. A stemmer was applied, and numbers and currencies removed. No stop words were used. We used manifestos from parties with relatively extreme positions on immigration (pro- and anti-) as reference texts (supplement S10). Two approaches were tested: using two reference texts, and using all parties in a given year as reference texts ($r=0.83$). Reference texts were set to the value in the pooled expert surveys, using election years close to 2003. Because of the often significantly shorter manifestos in the 1990s and because debates on immigration seem to evolve in the words used, we have not used reference texts from the 1990s or 2010s. The same two manifestos used as reference texts by Wordscores were used to indicate the direction of the underlying positions in Wordfish. All models were run with the default parameters. We used the raw scores and rescaled them empirically to fit the range 0 to 10 because neither the raw scores nor the ‘adjusted’ ones provided realistic estimates.⁷

There is no clear solution to infer party positions from manifestos when immigration is not mentioned at all, and we do not estimate party positions in this case (16 out of 283 manifestos, supplement S11). We explicitly do not draw on previous or subsequent positions, or on other issue domains like left-right positions, and we deliberately avoided searching for alternative

⁷ All estimates of party positions were re-scaled to a scale from 0 (pro-immigrant) to 10 (anti-immigrant) to ease comparison.

texts. Where parties do not mention immigration in their manifesto, we consider it likely that experts draw on left-right positions or other information and public statements as heuristics to infer (expected) positions on immigration where in reality there might be no (agreed) position. The same problem may apply to automated approaches if the texts are not restricted to sections known to be about immigration.

Silence can also be strategic: Anti-immigrant parties may mention immigration, whereas parties with less restrictive positions may say nothing or very little about immigration. This would be in line with expectations from salience (Budge et al., 1987; Rabinowitz and Macdonald, 1989) and issue ownership (Petrocik, 1996) theories, as parties are only expected to emphasize issues in which they can expect a competitive advantage (Bélanger and Meguid, 2008; Budge and Farlie, 1983; Green-Pedersen, 2007). Hence, we compare position estimates with a direct measure of salience: the proportion of words devoted to immigration in the manifesto. If salience reflects positions, longer sections on immigration in the manifesto will correlate with negative positions. For salience, we only consider whether *something* is written about immigration, not what. Because of its crudeness, we expect this approach to be only weakly associated with other approaches, but it serves as a useful baseline to judge what strong and weak associations between estimated party positions are.

To obtain a time series of expert data for the entire period, a range of expert surveys were pooled (Bakker et al., 2015; Benoit and Laver, 2006; Hooghe et al., 2010; Ladner et al., 2009; Lubbers et al., 2002; Rohrschneider and Whitefield, 2012). Multiple expert surveys need to be

aggregated because none of them cover the entire period and, because there is insufficient overlap between surveys to consider them separately – only the CHES is a time series. We only included surveys that asked experts to position parties on immigration ‘in general’ or ‘considering all aspects.’ The questions asked are comparable, but it was necessary to rescale the positions on a common scale. Because expert surveys often do not coincide with national elections, we have assessed three different methods to estimate the party position at the time of the election. First, we have considered a simple moving average over 7 years, covering the respective year plus three years before and after. The time span is necessary to bridge gaps in coverage and reduce the impact of individual surveys. Years before and after elections were included to emulate how experts are thought to place parties: combining evaluations of past performance and promised policies. Second, we have taken expert surveys that coincided with election years, and for elections or parties not covered, used the mean of the closest expert survey before and after. Third, we have calculated an exponential moving average over 7 years to give more weight to expert surveys that coincide with the election year than surveys further away in time. The different methods lead to almost identical estimates ($r=0.99$, supplement S12). In the following we use exponential moving averages to capture expert positions.

Associations between different methods

First we examine the associations between different methods at the level of individual manifestos. All countries and elections are pooled. Table 1 shows the Pearson and Spearman rank correlations for the different methods considered. For ease of interpretation, expert positions and sentence-by-sentence coding are used as reference categories. There are high correlations between expert positions, sentence-by-sentence coding, and the checklist approach. Moreover, Supplement S13 demonstrates that the different dimensions covered in the checklist approach plausibly capture a single policy dimension. For the automated approaches, the correlations are weaker. The correlations for the dictionary approach are somewhat stronger than for Wordscores and Wordfish, both of which surprisingly fare worse than the pure salience measure. The assumption that salience and (negative) positions will strongly correlate – stemming from salience and issue ownership theories – is not fully corroborated by the data: The correlation is at best moderate ($r=0.4-0.5$). The strongest associations are between expert positions, sentence-by-sentence coding, and the checklist approach: a high association between the positions estimated by manual coding and expert surveys – except for CMP coding ($r \leq 0.29$), which fares worse than the pure salience

measure.⁸ Automated coding measures are weakly associated to each other and weakly related with alternative measures.

Table 1: Correlations between Positions Obtained Using Different Methods

	Experts	Sentence	Checklist	Wordscores	Wordfish	Dictionary	Saliency	CMP
Experts	1.00	0.86	0.84	0.27	0.32	0.44	0.42	0.21
	(1.00)	(0.84)	(0.83)	(0.26)	(0.32)	(0.50)	(0.32)	(0.17)
Sentence-	0.86	1.00	0.85	0.38	0.33	0.44	0.51	0.36
by-Sentence	(0.84)	(1.00)	(0.83)	(0.35)	(0.35)	(0.48)	(0.42)	(0.27)

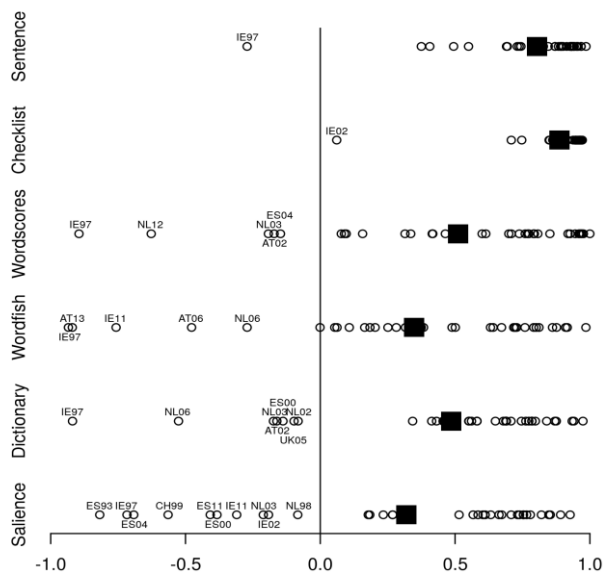
Notes: Pearson correlations, Spearman rank correlations in brackets (N=283 manifestos).

Figure 1 shows the correlation coefficients between two methods at a time, treating all manifestos per election as a case. Each circle represents a single election, while the black square is the average correlation coefficient. The fact that all black squares are on the right-hand side of the figure illustrates the positive correlations highlighted in Table 1. The individual circles illustrate the variance across elections. For the association between expert positions and checklist coding all observed correlations are positive. Even where these

⁸ See supplement S14 for correlations including derivative methods discussed above; supplement S15 demonstrates that the different methods all pick up positions on immigration using principal component analyses.

correlations are weak, they still tend to identify the same parties as pro- or anti-immigrant. There is less variability in the correlation coefficients for checklist coding than for sentence-by-sentence coding, and in both cases there is one outlying election. The few elections on the left-hand side of the figure indicate cases where the methods disagree on which parties are pro-immigrant and anti-immigrant. For many research applications, such unreliability in identifying the direction of the underlying concept is unacceptable, but in other situations the general tendency to get it right may suffice. Irish parties – with short sections on immigration – are problematic for all methods; all other cases affect automated methods.

Figure 1: Correlation Coefficients between Expert Positions and Other Methods

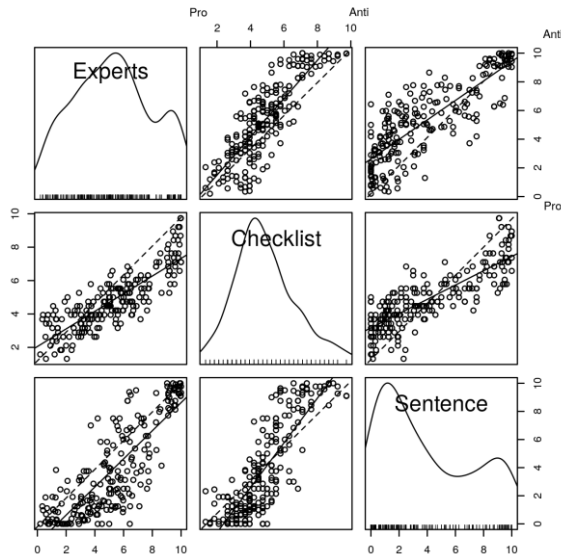


Notes: correlation coefficients are given as circles, with the black squares indicating the average correlation coefficient (N=43 elections, supplement S14 for numbers).

Identifying high correlations and a tendency to identify the same direction, however, is not sufficient to understand *how* the different methods compare. Depending on the substantive research question, it is important to anticipate how specific methods will behave. We show the association between expert positions, sentence-by-sentence coding, and the checklist approach graphically in Figure 2. Each circle represents a single manifesto, and we pool all countries and elections.⁹ The diagonal panels show kernel densities of the variables. Despite the relatively high correlations identified above, the distributions vary notably. For example, we note a hump at the extreme anti-immigrant end for expert positions not present in the checklist estimates. The sentence-by-sentence coding is bimodal, although with a dominant peak, and skewed towards pro-immigration positions. Relatively high correlations notwithstanding, the choice of method will have implications on the estimated party positions and their distribution.

⁹ Supplement S16 explores the relationship between Wordscores with different referencing methods and Wordfish graphically.

Figure 2: Associations between Experts, Checklist, and Sentence-by-Sentence Coding



Notes: Scatter plot matrix. OLS regression (straight) and equivalence (dotted) lines; diagonal panels show the density of each variable (N=283 manifestos).

If we examine the centre-left panel, we can see that checklists identify variation at the extreme anti-immigrant end, where experts do not. This might be a consequence of checklists drawing on specific sub-dimensions, which can increase the spread of values across parties. Extreme anti-immigrant positions aside, sentence-by-sentence coding tends to place parties somewhat more positive than experts, visible by the difference from the 45-degree line of equivalence. Sentence-by-sentence coding is also more likely to identify pro-immigrant positions than experts. This could be due to parties being less likely to express dislike of a

specific group of people directly ('person positivity' bias by parties), or because experts discount or give less weight to pro-immigrant positions from all parties (negative statements may 'stick' more: a 'cognitive' bias by experts). With the data at hand we cannot assess the process leading to these differences, and future research with experimental data might be illuminating.

The bottom-centre panel identifies an S-shaped relationship behind the strong correlation: Checklist and sentence-by-sentence coding tends to agree at the centre and both ends. Compared to checklist coding, for pro-immigrant parties, sentence-by-sentence coding tends to place parties as more pro-immigrant; for anti-immigrant parties, sentence-by-sentence coding tends to place parties as more anti-immigrant. These differences relate to how positions are identified. Expert positioning and sentence-by-sentence coding draw on whatever sub-dimensions the parties emphasize in public statements or manifestos; checklist coding assigns equal weight to all sub-dimensions that might not be given equal treatment by the parties. Consider a party with a restrictive position on irregular immigrants emphasized in the manifesto and public debates, and an expansive position on highly skilled immigrants mentioned in passing. Experts and sentence-by-sentence coding will identify an overall anti-immigrant position, whereas the two dimensions cancel each other out in the checklist approach, identifying a centrist position. The choice of approach will depend on the substantive research questions of the study.

Differences across countries and time

The patterns presented in Figure 1 and 2 may mask important differences across countries in how parties politicize immigration. In order not to conflate differences between political systems with differences between methods, we provide country-by-country analyses. The main results reported for the pooled analysis hold in all countries: The party positions obtained from experts, checklist, and sentence-by-sentence coding are consistently highly correlated (supplements S17, S18).

In Austria, Belgium, the Netherlands, and Switzerland, there are high correlations between expert positions and Wordscores for some years, but not for others, and sometimes they are negative. Correlations for Wordfish come with frequent sign changes. In these cases it is not apparent whether the different methods capture the same concept. The dictionary approach leads to relatively high associations in Belgium.

Spain and Ireland are of particular interest, because for several years immigration was less salient than in other countries, and manifestos dedicated much less space to immigration. Presumably as a result, we observe weaker correlations between different methods, although the strongest associations can still be found between experts, checklists, and sentence-by-sentence coding. Contrary to what we find in most other countries, the correlations between expert positions and checklist coding are higher than those between expert positions and sentence-by-sentence coding (except Ireland 2002).

In France, there are high correlations between expert positions and both sentence-by-sentence coding and the checklist approach. Expert positions and Wordscores correlate highly for the last three elections, but not for 1997. The United Kingdom is a special case, as all methods lead to similar estimates – except for the dictionary approach, which ‘fails’ for the 2005 elections.¹⁰ For other countries, we could not find stable associations for automated approaches, observing high correlations in some but not all elections.

Across the eight countries considered, checklists provide consistently high correlations with expert positions. By contrast, CMP positions tend to come with the lowest correlations with expert positions. In supplement S19 we demonstrate that the length of the texts on immigration is not the only factor driving the underperformance of automated methods. The correlations between expert positions and Wordscores tend to be stronger with longer manifestos – but they remain comparatively modest –, although not for Wordfish. The correlations of most methods increase with longer texts, leading to more consistent estimates. We refrain from suggesting a clear cut-off point when automatic approaches become viable, because this depends on the application and research question.

10 This suggests it may have been an ‘easy’ case for the testing of crowd-sourcing in Benoit et al. (2016).

Variance in the estimates

All estimates of party positions carry some uncertainty (McDonald and Budge, 2014). Here we examine standard deviations of the estimated party positions, acknowledging that there are other sources of error and bias. Automated methods have much smaller standard deviations than manual approaches,¹¹ while expert positions are somewhat in the middle (supplement S20). There is no apparent way to calculate standard deviations for the dictionary approach. Coder training can reduce the errors in manual coding, just like better instructions and survey design can reduce errors in expert surveys, but not to the extent that they would match automated coding. However, automated methods are not producing what would seem, at face value, ‘correct’ estimates of party positions.

The relatively large variation in estimated positions questions the idea that parties have precise or unequivocal positions on broad policy domains like immigration. Immigration is not a single issue, but a complex bundle of many sub-issues. Indeed, the variance discussed here can also be understood as the clarity of party positions (Lo et al., 2014). Some parties may benefit from clarity, namely parties that move to more extreme positions. For parties that move to the centre, by contrast, having a clear position can be an electoral disadvantage (Sommer-Topcu, 2015), especially in multidimensional competition spaces (Rovny, 2012, 2013).

¹¹ The standard deviations for checklist coding may be inflated because they also reflect heterogeneity in positions around the sub-issues measured.

With OLS regressions we examine under which circumstances greater variance is observed (Table 2). For sentence-by-sentence coding standard deviations become smaller the more extreme parties are, and for Wordscores and Wordfish as sections on immigration are longer.¹² All these associations are robust to the inclusion of country dummies and party-level controls (supplement S22).

Table 2: Standard Deviations, Difference from Median Position and Section Length

	Experts	Wordscores	Wordfish	Sentence-by-Sentence	Checklist
Constant	1.78	0.34	0.18	2.69	2.35
Difference from Median Position	0.01	0.01	-0.003	-0.33 ***	-0.004
	(0.15)	(0.01)	(0.004)	(0.05)	(0.03)
Section Length (Sentences)	-0.01	-0.003 ***	-0.002 ***	0.01 *	0.01 **
	(0.01)	(0.000)	(0.000)	(0.002)	(0.002)
N	34	197	217	215	218
Adjusted R ²	0.04	0.23	0.22	0.18	0.04

Notes: OLS regression with standard deviation of party position estimate as the dependent variable. Not all expert surveys report standard deviations or match election dates; hence the much smaller number of cases. For Wordscores two reference texts were used. * p<0.1, ** p<0.05, *** p<0.01

¹² Supplement S21 suggests that automated methods react to more data, whereas human coders react to clearer positions.

Discussion and conclusion

Issues relating to immigration have become increasingly important to structure political divides and party competition across established democracies, but our ability to examine how parties compete over immigration over time and across countries is hampered by the limited availability of suitable sources to estimate party positions. Using novel data for eight European countries, this article compared different methods that can be used to estimate positions on immigration from party manifestos and assessed how they fare relative to each other and compared to independent expert surveys. The data collected through the manual selection of sections dealing with immigration in party manifestos, allowed us to assess salience, as well as positions on an overall pro- vs anti-immigrant continuum and on a number of sub-issues relating to immigration and integration policies. This approach produces a rich array of data that will help researchers study party positions on immigration in nuanced ways and with various models of party competition in mind.

While the focus on immigration responds to a substantive interest, the analyses speak to common challenges in the estimation of party positions on policy dimensions that do not obviously or always overlap with the dominant party competition divide – commonly left-right – such as positions on the European Union, or on the environment. Politicization of the specific policy area may be uneven across time, countries, and parties, and the sections dedicated to the issue on party manifestos may be sometimes relatively short. Moreover, the rapid changes in the debate combined with constant framing and reframing processes of an

emerging issue may be particularly challenging for automated approaches that rely on a stable relationship between words and positions (Grimmer and Stewart, 2013). Our findings suggest that all methods are feasible and pick up the same underlying dimension, but they produce clearly different estimations, especially at the extreme ends, and not all produce what seem valid estimates at face value.

Our evaluation of the various methods departed from an agnostic position as to which method might be more suitable for the estimation of party positions on immigration. By comparing results with positions derived from expert surveys, manifestos are confirmed as an excellent source for positioning parties, especially because it is possible to go back in time and explore the emergence of new issues as they arise. Expert positions and the various methods of manual coding lead to very similar results in nearly all cases, and for some countries and elections automated approaches come close, too. The automated approaches, however, have unexpected outliers, despite a reasonable ideological dominance assumption (Grimmer and Stewart, 2013) given that sections on immigration were manually chosen. We find that scarce data – short sections on immigration – negatively affect placing parties, although Wordscores appears to be particularly affected. Despite the heavy input of expert knowledge during dictionary creation, the automated dictionary approach still underperformed compared to expert surveys and manual coding. This article also explored the nature of associations between methods, and there is no single ‘best’ method suitable for all research

Our analysis highlights that using the historical codes of the CMP to capture party positions on immigration leads to poor estimations and considerable errors of placement. In many cases, CMP coding performed worse than a simple salience approach, and we recommend researchers not to rely on the CMP codes for these purposes with data older than the 2014 5th edition of the codebook, where specific sub-codes were introduced.

While automated methods come with the benefits of low cost and short data processing timescales, they have the disadvantage of seemingly erroneous estimations of positions for complex and specific policy areas, especially (but not only) when the textual data are scarce. By contrast, manual coding methods (and expert surveys) are much more time consuming, costly and less reliable, but lead to more valid estimates. The analyses in this article are designed to help other researchers choose the most appropriate method for their particular applications.

A key contribution of this article is to show that manual coding of manifestos as a whole – the ‘checklist’ approach – turns out to be an efficient and valid method to code large numbers of manifestos in a relatively short time: typically between 10 and 20 percent of the time needed for manual sentence-by-sentence coding. In research on complex policy domains not necessarily overlapping with the left-right divide, we expect that often the checklist approach provides a good balance: it is much quicker than sentence-by-sentence coding, yields comparable results, and for shorter corpora may even be superior. Crowd-sourcing could

perhaps make the checklist approach even more cost-effective, but we suspect the task is too complex and the texts too long to work in such an environment.

The checklist approach does not seem to produce many mis-classifications of party positions, and provide a more nuanced measurement of the different positions at the extremes of the continuum, which expert surveys and sentence-by-sentence coding have more difficulty in capturing. Checklist coding was robust across a considerable number of countries and parties with varying degrees of salience of the policy domain analysed. It may be the least common of the approaches, but our results suggest that this is undeserved, as it combines relative cost-efficiency without obvious loss in validity and reliability. Although preparing the checklist items requires some work and specialist knowledge, it is no more laborious than preparing a full codebook for sentence-by-sentence manual coding or preparing the questionnaire of an expert survey; and it takes considerably less time to undertake data collection.

Silence is a challenge that deserves separate investigation: When parties choose not to make statements in a particular domain, placing them is both conceptually and empirically difficult. Further research should also pay close attention to the variance in party positions, because positioning parties is fraught with uncertainties. With any method used, the uncertainty about the positions is non-negligible. This is probably both a reflection of the imperfection of our methods and the fact that parties are often not clear and consistent in all the positions they hold on any given issue or policy field.

In conclusion, with different methods to estimate party positions from manifestos, it is important that researchers make an informed choice (Laver, 2014). In addition to offering a solution for a stable and sustainable way to measuring party positions on immigration over time, we hope that this assessment will help other researchers choose the most appropriate approach for their substantive research questions when dealing with other narrowly defined (but sometimes complex) issues or policy fields.

References

- Bakker R, Vries C de, Edwards E, et al. (2015) Measuring party positions in Europe The Chapel Hill expert survey trend file, 1999–2010. *Party Politics* 21(1): 143–152.
- Bélanger É and Meguid BM (2008) Issue salience, issue ownership, and issue-based vote choice. *Electoral Studies* 27(3): 477–491.
- Benoit K and Laver M (2006) *Party Policy in Modern Democracies*. London: Routledge.
- Benoit K and Laver M (2007) Estimating party policy positions: Comparing expert surveys and hand-coded content analysis. *Electoral Studies* 26(1): 90–107.
- Benoit K, Conway D, Lauderdale BE, et al. (2016) Crowd-sourced Text Analysis: Reproducible and Agile Production of Political Data. *American Political Science Review* 110(2): 278–295.
- Black, D. (1948) On the Rationale of Group Decision-Making. *Journal of Political Economy* 56 (1): 23–34.
- Budge I (2000) Expert judgements of party policy positions: uses and limitations in political research. *European Journal of Political Research* 37(1): 103–113.
- Budge I and Farlie DJ (1983) *Explaining and predicting elections: Issue effects and party strategies in twenty-three democracies*. George Allen & Unwin London.
- Budge I, Robertson D and Hearl D (1987) *Ideology, strategy, and party change: spatial analyses of post-war election programmes in 19 democracies*. Cambridge University Press.
- Budge I, Klingemann H, Volkens A, et al. (2001) *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments, 1945-1998*. Oxford: Oxford University Press.
- Castles F and Mair P (1984) Left–right political scales: some ‘expert’ judgments. *European Journal of Political Research* 12: 73–88.
- Däubler T, Benoit K, Mikhaylov S, et al. (2012) Natural sentences as valid units for coded political texts. *British Journal of Political Science*.

- Dinas E and Gemenis K (2010) Measuring Parties' Ideological Positions With Manifesto Data A Critical Evaluation of the Competing Methods. *Party Politics* 16(4): 427–450.
- Enelow JM and Hinich MJ (1984) *The Spatial Theory of Voting: An Introduction*: Cambridge University Press.
- Green-Pedersen C (2007) The growing importance of issue competition: The changing nature of party competition in Western Europe. *Political Studies* 55(3): 607–628.
- Grimmer J and Stewart BM (2013) Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis* 21(3): 267–297.
- Gudbrandsen F (2010) Partisan influence on immigration: the case of Norway. *Scandinavian Political Studies* 33(3): 248–270.
- Harmel R, Janda K and Tan A (1995) Substance vs. Packaging: An Empirical Analysis of Parties' Issue Profiles. In: *Annual Meeting of the American Political Science Association*. Available from: <http://www.janda.org/bio/parties/papers/Harmel%20et%20al,%201995.pdf> (accessed 11 October 2012).
- Hooghe L, Bakker R, Brigeovich A, et al. (2010) Reliability and validity of the 2002 and 2006 Chapel Hill expert surveys on party positioning. *European Journal of Political Research* 49(5): 687–703.
- Huber J and Inglehart R (1995) Expert interpretations of party space and party locations in 42 societies. *Party Politics* 1(1): 73–111.
- Katz RS and Mair P (1992) Introduction: The Cross-National Study of Party Organizations. In *Party Organizations: A Data Handbook on Party Organizations in Western Democracies 1960-90*, eds. Richard Katz and Peter Mair. London: Sage. 1-20.
- Katz RS and Mair P (1994) The Evolution of Party Organizations in Europe: The Three Faces of Party Organization. *American Review of Politics* 14: 593-617.
- Klüver H (2009) Measuring Interest Group Influence Using Quantitative Text Analysis. *European Union Politics* 10(4): 535–549.
- Kriesi H, Grande E, Lachat R, et al. (2006) Globalization and the transformation of the national political space: Six European countries compared. *European Journal of Political Research* 45(6): 921–956.

- Kriesi H, Grande E, Dolezal M, et al. (2012) *Political Conflict in Western Europe*. Cambridge: Cambridge University Press.
- Ladner A, Schwarz D and Fivaz J (2009) Die Schweizer Parteien im politischen Raum - eine Analyse der politischen Positionen ihrer Kandidierenden bei den Nationalratswahlen 2007. *Working paper de l'IDHEAP* 2009(1).
- Laver M (2014) Measuring Policy Positions in Political Space. *Annual Review of Political Science* 17(1): 207–223.
- Laver M, Benoit K and Garry J (2003) Extracting policy positions from political texts using words as data. *American Political Science Review* 97(2): 311–331.
- Lo J, Proksch SO and Slapin JB (2014) Ideological Clarity in Multi-Party Competition: A New Measure and Test Using Election Manifestos. *Working Paper*.
- Lowe W (2008) Understanding Wordscores. *Political Analysis* 16(4): 356.
- Lowe W and Benoit K (2013) Validating estimates of latent traits from textual data using human judgment as a benchmark. *Political Analysis* 21(3): 298–313.
- Lubbers M, Gijsberts M and Scheepers P (2002) Extreme right-wing voting in Western Europe. *European Journal of Political Research* 41(3): 345–378.
- Marks G, Hooghe L, Steenbergen M, et al. (2007) Crossvalidating data on party positioning on European integration. *Electoral Studies* 26(1): 23–38.
- McDonald MD and Budge I (2014) Getting it (approximately) right (and center and left!): Reliability and uncertainty estimates for the comparative manifesto data. *Electoral Studies* 35: 67–77.
- McDonald MD, Mendes S and Kim M (2007) Cross-temporal and cross-national comparisons of party left-right positions. *Electoral Studies* 26(1): 62–75.
- Odmalm P (2012) Party competition and positions on immigration: Strategic advantages and spatial locations. *Comparative European Politics* 10(1): 1–22.
- Petrocik JR (1996) Issue ownership in presidential elections, with a 1980 case study. *American Journal of Political Science* 40(3): 825–850.
- Rabinowitz G and Macdonald SE (1989) A directional theory of issue voting. *The American Political Science Review*: 93–121.

- Ray L (2007) Validity of measured party positions on European integration: Assumptions, approaches, and a comparison of alternative measures. *Electoral Studies* 26(1): 11–22.
- Revelle W (2015) *psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois: Northwestern University. Available from: <http://CRAN.R-project.org/package=psych>.
- Rohrschneider R and Whitefield S (2012) *The Strain of Representation: How Parties Represent Diverse Voters in Western and Eastern Europe*. Oxford University Press.
- Rooduijn M and Pauwels T (2011) Measuring Populism: A Comparison of Two Methods of Content Analysis. *West European Politics* 34(6).
- Rovny J (2012) Who emphasizes and who blurs? Party strategies in multidimensional competition. *European Union Politics* 13(2): 269–292.
- Rovny J (2013) Where do radical right parties stand? Position blurring in multidimensional competition. *European Political Science Review* 5(1): 1–26.
- Ruedin D (2013) Obtaining Party Positions on Immigration in Switzerland: Comparing Different Methods. *Swiss Political Science Review* 19(1): 84–105.
- Somer-Topcu Z (2015) Everything to Everyone: The Electoral Consequences of the Broad-Appeal Strategy in Europe. *American Journal of Political Science* 59(4): 841–854.
- Steenbergen MR and Marks G (2007) Evaluating expert judgments. *European Journal of Political Research* 46(3): 347–366.
- Van der Brug W, D’Amato G, Berkhout J, et al. (eds) (2015) *The Politicisation of Migration*. Abingdon: Routledge.